

2

AEOSR-TR- 90 0782

AD-A224 558

## ROLE OF RETINOCORTICAL PROCESSING IN SPATIAL VISION

DTIC FILE COPY

Final Report

June 1990

By: D. H. Kelly, Principal Scientist  
Sensory Sciences Research Laboratory

Prepared for:

United States Air Force  
Air Force Office of Scientific Research  
Directorate of Life Sciences  
Building 410  
Bolling Air Force Base  
Washington, D. C. 20332-6448

Attn: Dr. John F. Tangney

Contract F49620-87-K-0009

SRI Project 3558

SRI International  
333 Ravenswood Avenue  
Menlo Park, California 94025-3493  
(415) 326-6200  
Telex: 334486

DTIC  
ELECTE  
JUL 26 1990  
S E D



REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 1990 June	3. REPORT TYPE AND DATES COVERED Final Report 890501 to 900501		
4. TITLE AND SUBTITLE Role of Retinocortical Processing in Spatial Vision		5. FUNDING NUMBERS F49620-87-K-0009 61102F 2313 A5		
6. AUTHOR(S) Kelly, Donald H.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) SRI International 333 Ravenswood Avenue Menlo Park, CA 94025		8. PERFORMING ORGANIZATION REPORT NUMBER 3558 Final Report		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research Directorate of Life Sciences Building 410 Bolling AFB, DC 20332-6448		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words) <p>This project was an attempt to understand the major functions of early vision by considering how various components cooperate in preparing visual information to organize our perceptions of the world around us. To this end, and with the cooperation of SRI's Machine Vision Group, we assembled some of these functions in a computational working model, which can graphically display the spatial structure of the information at a given stage of the visual process, in the form of a two-dimensional intensity array (or "image"). This capability was developed to facilitate the study and comparison of retinal and cortical inputs and outputs of spatial information.</p> <p>The individual components of the model are well known, and the relations among them are based on available data from the literature. However, two aspects of this project seem novel. One is the exploitation of powerful, state-of-the-art tools of computational vision, such as Symbolics 3600-series LISP machines, to create and display our results. These tools were developed primarily for artificial intelligence purposes; they have rarely been used for basic studies in human vision. (JRS)</p>				
14. SUBJECT TERMS Spatial vision; retinocortical projection; computational model		15. NUMBER OF PAGES 31		
		16. PRICE CODE		
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED	

**SECURITY CLASSIFICATION OF THIS PAGE**

**CLASSIFIED BY:**

**DECLASSIFY ON:**

- Inhomogeneous filtering by retinal receptive fields
- Remapping of visual space by the retinocortical projection
- Image analysis by receptive fields of the striate cortex
- Multiple fixations of a single scene

A number of important aspects of early vision (e.g., depth, motion, and chromatic responses) were intentionally omitted in order to keep the model of manageable proportions during this first-phase development. The most important one to be treated in any follow-on phase would be the chromatic aspects, for reasons that have emerged from very recent work on spatiochromatic cortical modeling.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification _____	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



## CONTENTS

I	RESEARCH OBJECTIVES .....	1
II	PROJECT RESULTS THROUGH APRIL 1990.....	3
	A. Relation Between Retinal and Cortical Filtering .....	3
	B. Orientation Selectivity.....	7
	C. Computational Strategy.....	9
	D. Cortical Output Images .....	9
	E. The Stable Frame .....	18
III	SIGNIFICANT MILESTONES.....	20
IV	RECOMMENDATIONS FOR FURTHER WORK.....	21
V	LIST OF RESEARCH PERSONNEL .....	22
VI	INTERACTIONS WITH SCIENTIFIC COMMUNITY.....	23
	REFERENCES.....	24

## LIST OF ILLUSTRATIONS

1	Outline of the Model and Chronology of Work.....	2
2	Cortical Projections of Homogeneous, Isotropic White Noise Filtered by (a) In-homogeneous Retinal Receptive Fields, and (b) Homogeneous Gabor Process.....	4
3	Comparison of Cortical Input and Output Computations Discussed in Text.....	6
4	Ratio of retinocortical magnification (according to the Schwartz formula) to retinal receptive field size, as used in the model.....	6
5	Cortical projection of building scene, with (a) no filtering, (b) linear, homogeneous Gabor filter, and (c) six-orientation, "winner-take-all" line detection .....	8
6	Winner-take-all filtering algorithm with cortical projection of white noise; the Gabor frequencies used in (a) and (b) were two octaves apart .....	10
7	Cortical projection of conference room scene, fixated on man's spectacles; winner-take-all algorithm at four Gabor frequencies, each separated from its neighbors by one octave.....	11
8	Same as Figure 7 but fixated on woman's left eye .....	13
9	Same as Figures 7 and 8 but fixated on center of notebook.....	14
10	Cortical projection of "Lena" portrait, fixated on bridge of her nose. Same computation at same frequencies as Figures 7-9.....	15
11	Same as Figure 10 but fixated on left brim of model's hat.....	16
12	Same as Figures 10 and 11 but fixated on upper right brim of model's hat.....	17
13	Cortical projections for three fixations of conference room scene at the same, intermediate Gabor frequency; somehow the information in several projections like this must be integrated to form a stable percept of the scene.....	19

## I RESEARCH OBJECTIVES

Our goal was to develop a computational model of the "front-end" stages of human spatial vision, including the retina, retinocortical pathways, and primary visual cortex (V1), as illustrated schematically in Figure 1. This computational product was to be a functional, *working* model, which processes the entire stimulus pattern by appropriate algorithms, and can depict its representation at each stage in graphic imagery.

To make this task more manageable, important but noncritical simplifications were made. The model was confined to monocular, photopic, achromatic, quasi-stationary vision. Motion was considered only to the extent that normal spatial processing requires minimal eye movements. Binocularity was considered only by constraining V1 to leave room for interleaved right- and left-eye connections.

Important parts of this complex system have been modeled in other studies. Our main goal was to try to make them all fit together. In doing that, we expected to encounter problems that had not shown up before, and we have. In the course of trying to solve them, much has been learned, as described in the next section.

The chronology of our efforts is indicated approximately by the vertical arrows in Figure 1. Our goals of simulating the retinocortical projection and integrating it with inhomogeneous retinal filtering were achieved during the first two years of the project, as described in detail in our Annual Reports 1 and 2. During the third year, we attempted to incorporate into our model the spatial-vision aspects of postsynaptic processing in the visual cortex.

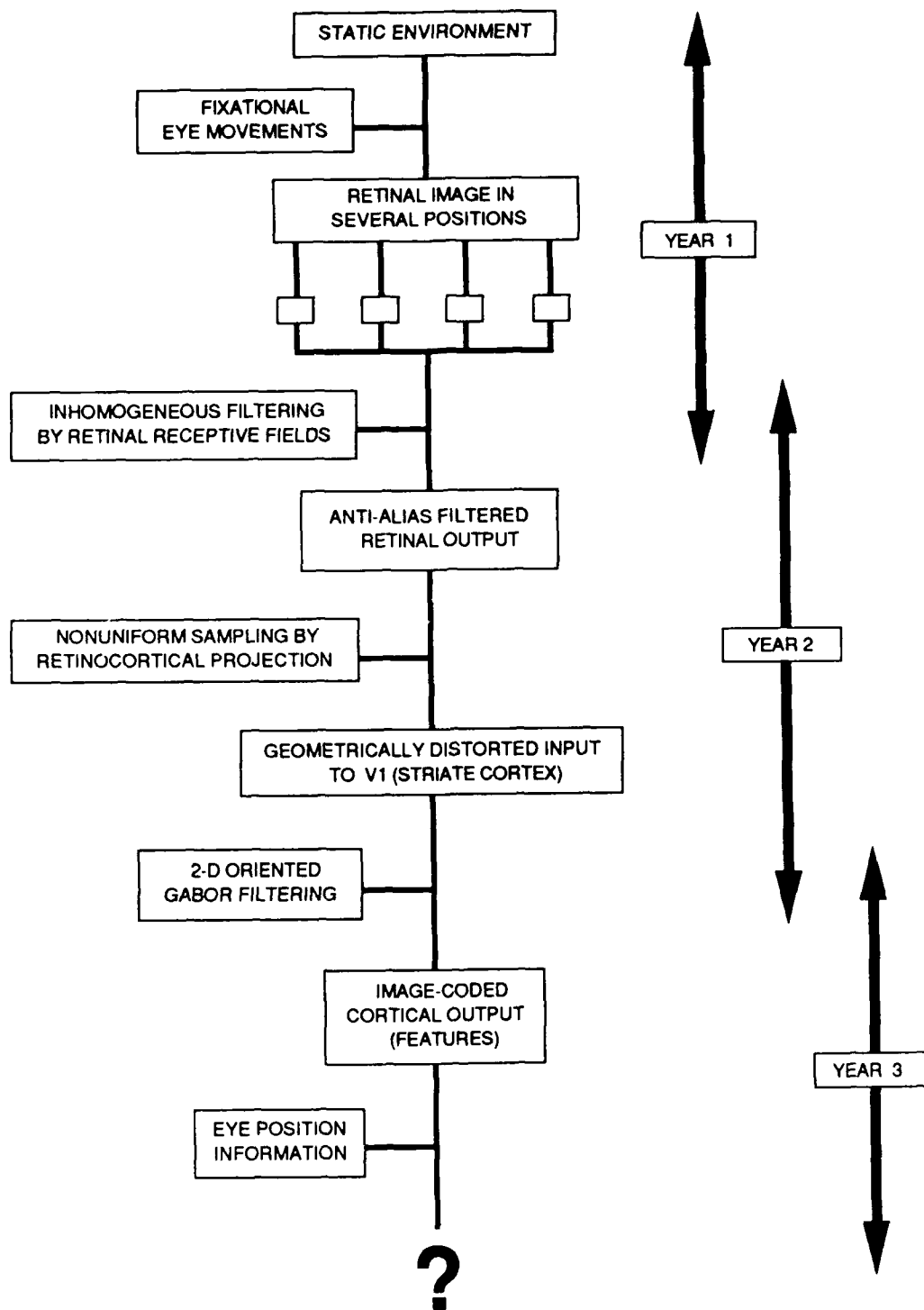


FIGURE 1 OUTLINE OF THE MODEL AND CHRONOLOGY OF WORK

## II PROJECT RESULTS THROUGH APRIL 1990

### A. RELATION BETWEEN RETINAL AND CORTICAL FILTERING

The transformation from retinal image to cortical input involves two important functions of eccentricity in the visual field: (1) the variation in ganglion-cell receptive-field size (retinal inhomogeneity), and (2) the variable rate of spatial sampling (cortical magnification) by which these retinal cells are connected to the striate cortex (V1). Since the two functions do not track each other perfectly (see Figure 1 of Annual Report 2), our model was originally designed to represent the effects of each independently. Thus, it can create accurately filtered and distorted cortical inputs [like those shown in Figures 4, 6(b), 7, and 8(b) of Annual Report 2].

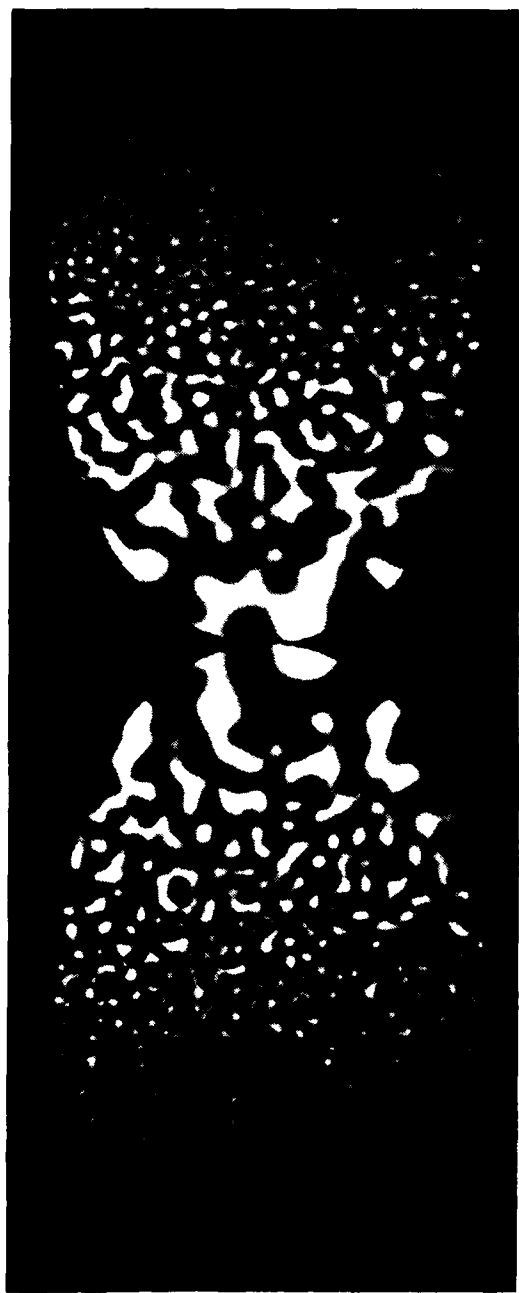
As the final phase of this project, we undertook to model the postsynaptic cortical processing of those inputs. Cortical filtering differs from retinal filtering in two important ways, both of which we simulated. First, cortical filtering is much more homogeneous. It has been said that the layers of the primary visual cortex seem almost crystalline in their regularity. This permits a useful shortcut in the filter computation.

Second, cortical filtering is not isotropic but orientation selective. The spatial weighting functions we use to simulate cortical filtering are patterned after the receptive fields of Hubel and Weisel's line-detector cells. These cortical receptive fields are oblong in shape, with a responsive bar in the center, flanked by two parallel antagonistic bars. There are many ways to create such a kernel, but the best known, usually called a two-dimensional Gabor function, is merely a spatial sinusoid multiplied by an elliptical Gaussian envelope. Gabor functions are mathematically simple, and their similarity to cortical receptive fields is adequate for our purposes. They come in many forms, depending on the frequency, orientation, and phase of the sinusoidal component, and on the ellipticity of the envelope. We use only even functions with 2-to-1 ellipticity [as shown in Figure 3(b) of Annual Report 2], at a small number of frequencies and orientations.

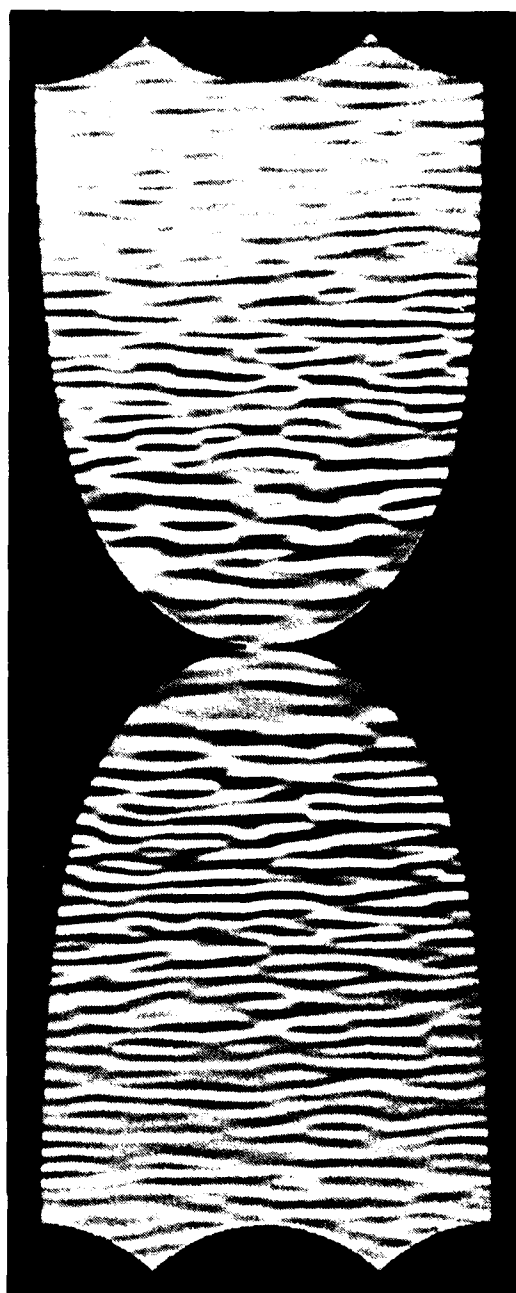
The contrast between isotropic but inhomogeneous filtering at the retina and homogeneous but anisotropic filtering at the cortex is illustrated by the responses of these two processes to a homogeneous, isotropic, white-noise test pattern, as shown in Figure 2. White-noise images were used to test our model throughout, particularly the cortical responses described below.

Because we assume cortical homogeneity, the filtering at this stage can be performed in the spatial frequency domain, taking the two-dimensional Fourier transform of the filter kernel, multiplying it by the transform of the input, and inverse-transforming this product to obtain the output image. Computationally, that is much faster than the cumbersome convolutions we had to use in the retina of our model.





(a)



(b)

FIGURE 2 CORTICAL PROJECTIONS OF HOMOGENEOUS, ISOTROPIC WHITE NOISE  
FILTERED BY (a) INHOMOGENEOUS RETINAL RECEPTIVE FIELDS, AND  
(b) HOMOGENEOUS GABOR PROCESS

But it works correctly only if we assume that retinal inhomogeneity is exactly compensated by retinocortical magnification. This assumption is needed because of the way cortical receptive fields are measured. The receptive fields of visual cells are not defined in terms of the signal at any preceding synapse. They are always measured with respect to the optical stimulus in the visual field. Thus, even if the receptive field of a given cortical cell is a perfect Gabor function, that function is not just the kernel of a filtering process located in the cortex. It represents the entire sequence of all spatial filtering processes from the retina through the cortex.

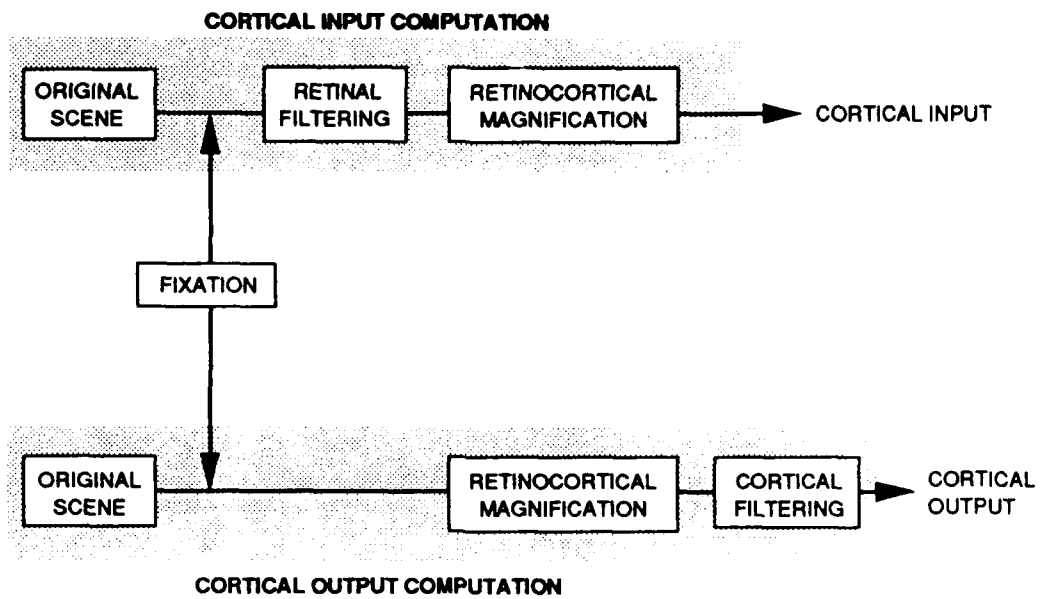
One way to treat our retinal filtering process as part of that sequence would be to deconvolve a virtual retinal filter from the cortical measurements, and then apply what was left to the cortical input. Since one is homogeneous and the other is not, that would be a very complicated process with little reward (because we already have the overall system kernel).

Fortunately, there is a much simpler way to accomplish almost the same thing. Our model has been so constructed that retinal filtering can be eliminated while the retinocortical magnification is left intact. [For a distorted but unfiltered image of this type, see Figure 2(b) of Annual Report 2.] Except for two minor defects, this is the type of input needed for our Gabor filters.

We can cure the first of these problems: If we carried out the full deconvolution procedure, then the cortical stage of our model wouldn't have to deal with the dc component of the original scene, because the Laplacian filters we use at the retinal stage have no dc response. But in the computation illustrated here at the bottom of Figure 3, where retinal and cortical filtering are combined as a Gabor kernel, we must remove the dc component in the only way available, by notching the spectrum of our even Gabor filters. (In the space domain, these modified Gabor functions are very similar to the originals.)

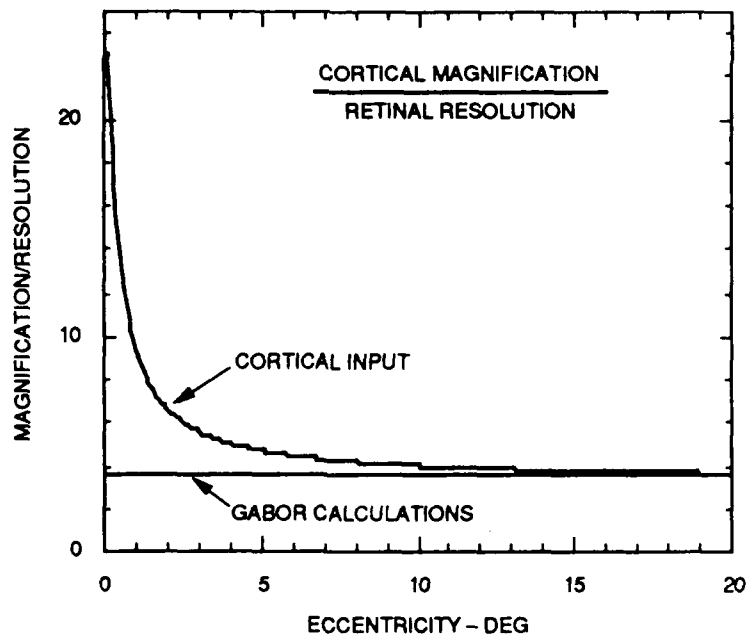
The second problem is more fundamental: When we model retinocortical information transfer correctly, the cortical magnification increases more rapidly near the fovea than the retinal resolution does (as shown in Figure 1 of Annual Report 2). The ratio of those two quantities in our cortical input calculations is shown here in Figure 4. If the retinocortical magnification exactly compensated for the inhomogeneity of the retina, the curve would be flat, and it almost is, outside the parafovea. But this scale invariance doesn't hold close to the fovea, where the cortical magnification is too great, even for the very fine sampling of foveal receptive fields. This effect was taken into account in calculating our cortical input images.

However, we can greatly simplify the simulation of Gabor filtering at the cortex by using remapped but unfiltered cortical inputs. That forces us to assume that the size of the retinal filter kernel is inversely proportional to the retinocortical magnification everywhere, as represented by the horizontal line in Figure 4. That is not the correct relation, but since it is a good approximation over most of the visual field, we used it in our shortcut computation of cortical outputs. We also checked those outputs by comparing them with the (correctly calculated) cortical inputs, particularly in the foveal region. The only other practical solution (described in Section IIC.2 of Annual Report 2) would have confined our results to much too narrow a range of spatial frequencies.



ke-060190-jd

FIGURE 3 COMPARISON OF CORTICAL INPUT AND OUTPUT COMPUTATIONS DISCUSSED IN TEXT



ke-060290-jd

FIGURE 4 RATIO OF RETINOCORTICAL MAGNIFICATION (according to the Schwartz formula) TO RETINAL RECEPTIVE FIELD SIZE, AS USED IN THE MODEL

## B. ORIENTATION SELECTIVITY

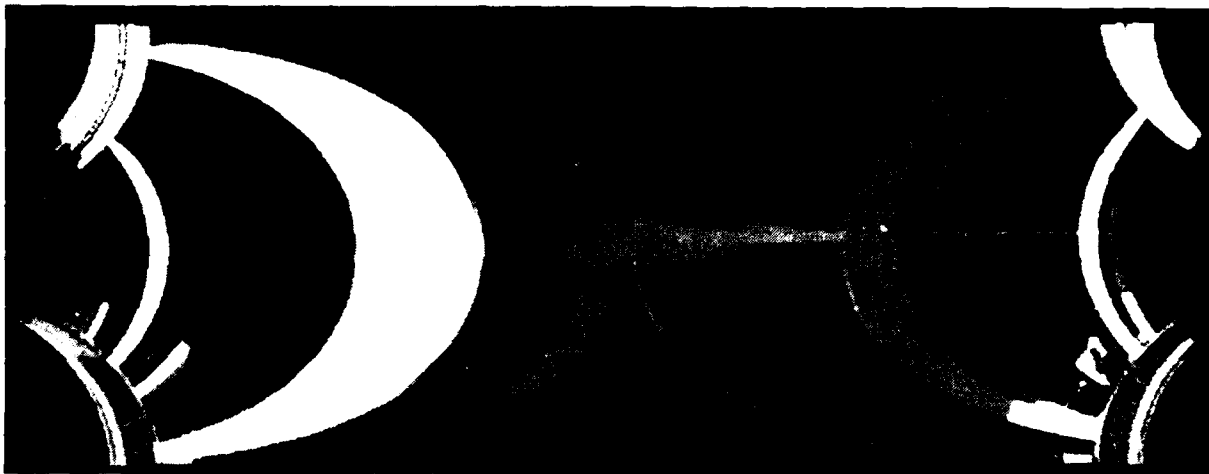
In Figure 2(b) we applied a Gabor filter to the (Schwartz) cortical projection of the noise target--a perfectly linear, spatial filtering process. The resulting tiger stripes merely show that the filter and the test pattern are doing their jobs properly: a line detector will detect lines, if there are any, and random noise contains lots of lines. The filter doesn't simply create them--it lifts them out of the noise, by correlation. Obviously this Gabor filter was oriented vertically. Other orientations would reorient the tiger stripes accordingly.

But a fixed orientation in cortical coordinates doesn't seem to correspond to anything significant, other than meridians or circles in the visual field, and hence would not be a very useful kind of output. To arrive at a more useful (and more physiological) scheme, first consider the projected images of the building shown in Figure 5. This scene also served as a test target, because it contains many horizontal and vertical, straight lines. But in the cortical projection shown in Figure 5(a), all the lines except the horizontal meridian are curved. A linear, vertical Gabor filter can only pick out vertical segments of these lines where they occur, as illustrated in Figure 5(b). We need a more sophisticated line detector--one that responds to each of these lines as far as it runs, regardless of orientation.

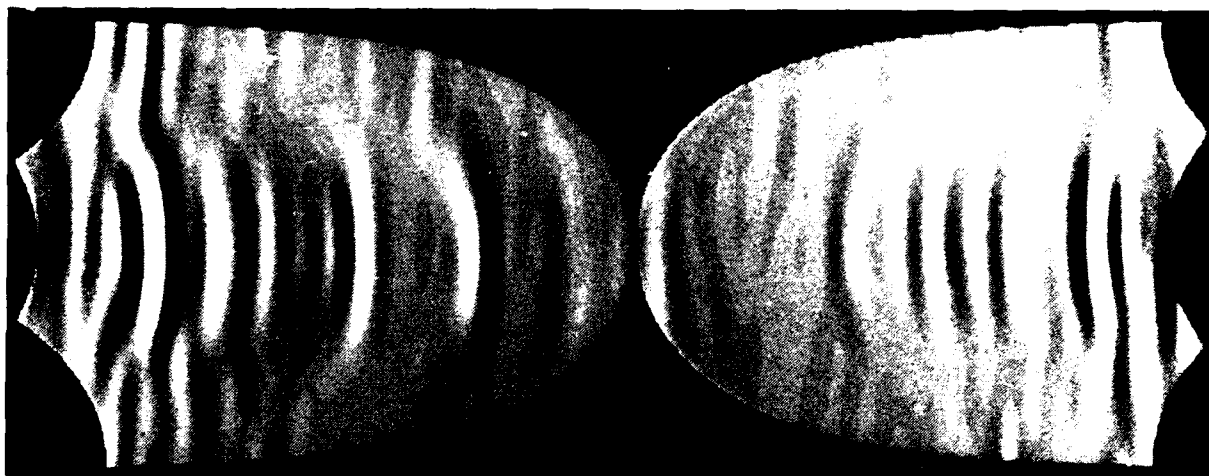
An algorithm that does this could be created in the following way. Suppose we rotate a Gabor filter of given frequency about a fixed point in the cortical image until we find its strongest response, and store that strength and orientation. Then, we move to the next point and repeat the procedure, continuing until we have a complete output of Gabor responses at that frequency, each at the optimum orientation for its location, suppressing all other orientations. Obviously, that constitutes a very nonlinear filtering process.

The result of applying such an algorithm to the building scene is shown in Figure 5(c), for the same (relatively low) frequency of Gabor function used in Figure 5(b). (Other spatial frequencies, covering a broad range, are used for other illustrations in what follows.) This winner-take-all scheme is similar to image-coding techniques that have been used by Watson, Daugman, and others.

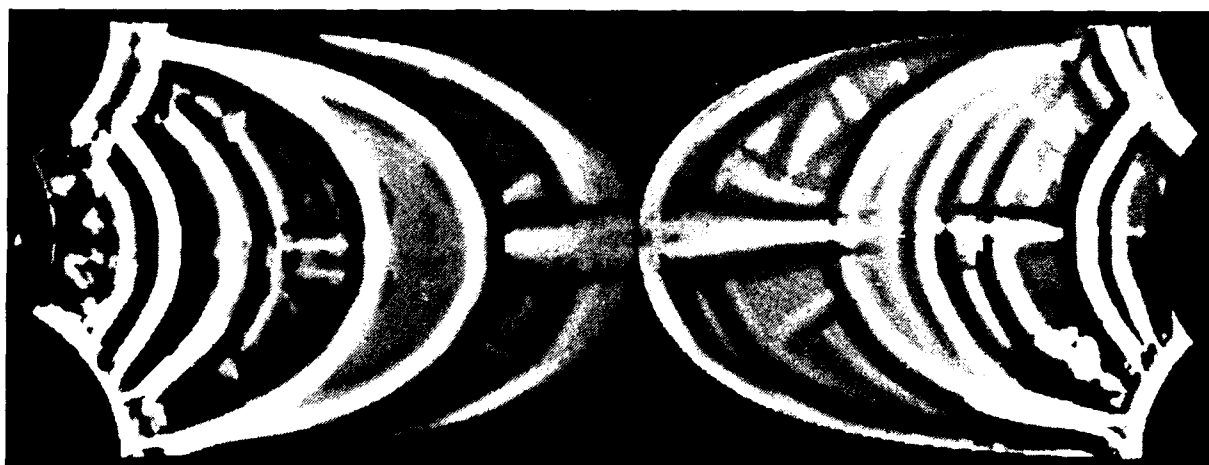
Physiologically, this scheme corresponds to the fact that a strongly stimulated line-detector cell may inhibit cells of other orientations for the same location in the visual field. The orientation bandwidth of these cells is about 30 degrees, like that of our Gabor filters. This suggests that instead of a continuous scan of orientation, we need only about six different orientations, spaced like the hours on a clock face. [Note that Figure 5(c) captures the smoothness of the curves in Figure 5(a); segments are scarcely visible.] This orientation spacing was therefore used for all our cortical simulations.



(a)



(b)



(c)

FIGURE 5 CORTICAL PROJECTION OF BUILDING SCENE, WITH (a) NO FILTERING, (b) LINEAR, HOMOGENEOUS GABOR FILTER, AND (c) SIX-ORIENTATION, "WINNER-TAKE-ALL" LINE DETECTION

### C. COMPUTATIONAL STRATEGY

The way we actually made these multiorientation images sounds very different from the procedure described above. In fact, the result is exactly the same, but computing time and storage requirements are much less. This improved algorithm can be illustrated with white-noise targets. First we compute one orientation in the usual way, convolving the desired Gabor function with the cortical input to obtain, for example, the image shown in Figure 2(b); this image is stored. Then we compute the next orientation, but this second image is not stored.

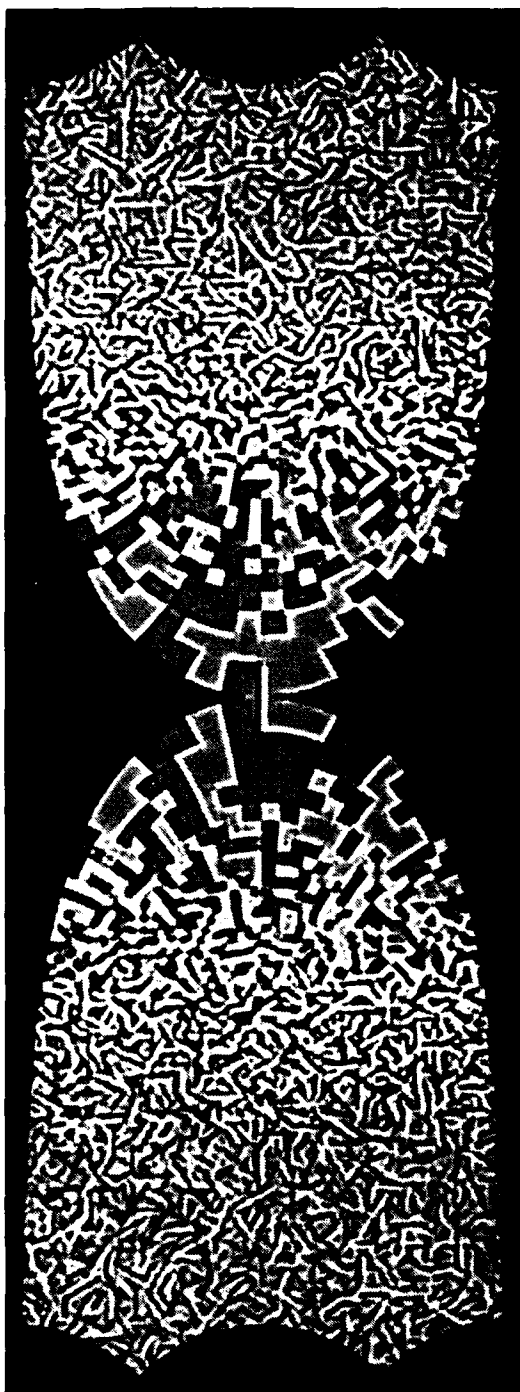
Instead, as each new output pixel is obtained, we compare the square of its value with the squared value of the same pixel in the stored image. (Any measure of its contrast, or distance from the zero mean, would do, but squaring is fast.) If the new value is greater, then the old pixel is replaced by the new one; if not, it remains. When this has been done for all pixels, the stored result is a two-orientation, winner-take-all image. If we reorient the kernel and repeat the process, the stored result becomes a three-orientation, winner-take-all image, and so on. These cortical images use up a great deal of memory, so it is an important advantage to store only one, instead of several. With six orientations 30 degrees apart (our standard procedure), a complete output image can be calculated in about 45 minutes.

White-noise tests are shown in Figure 6 for two Gabor frequencies. The white-noise input is exactly the same in Figures 6(a) and 6(b)--only the Gabor filter has been changed. We tested all our Gabor frequencies and bandwidths with white noise as in Figure 6, but for this report we will illustrate those properties with more realistic inputs.

### D. CORTICAL OUTPUT IMAGES

At this point, we return to the conference-room scene (Figures 2 and 4 of Annual Report 2), with the fixation point centered between the man's eyes. The four spatial frequencies shown in Figure 7 are our standard Gabor frequencies, with the lowest frequency at the top of the figure. Moving downward, spatial frequency doubles from each image to the next, so the total frequency range shown here is 8 to 1 (three octaves). Note, however, that the appropriate units for these Gabor frequencies would be cycles per millimeter of cortical surface. (They cannot be uniquely specified in terms of real-world cycles per degree of visual angle, because that varies with retinocortical magnification.)

Other fixations and other scenes generally share the properties illustrated here. Not surprisingly, the greatest similarity to our cortical input images (Figures 4 through 16 of Annual Report 2) occurs at the intermediate frequencies, which are closest to the peak frequency of the retinally filtered input. There is very little overlap between adjacent frequency bands in Figure 7 et seq., yet many elements of the scene can be identified in all four images. On the other hand, the images are almost as different as four artist's styles, from abstract painting for the lowest frequency to pen-and-ink sketch for the highest. It would be easy to believe that they represent four quite different types of information about the visual environment.



(a)



(b)

FIGURE 6 WINNER-TAKE-ALL FILTERING ALGORITHM WITH CORTICAL PROJECTION OF WHITE NOISE; THE GABOR FREQUENCIES USED IN (a) AND (b) WERE TWO OCTAVES APART

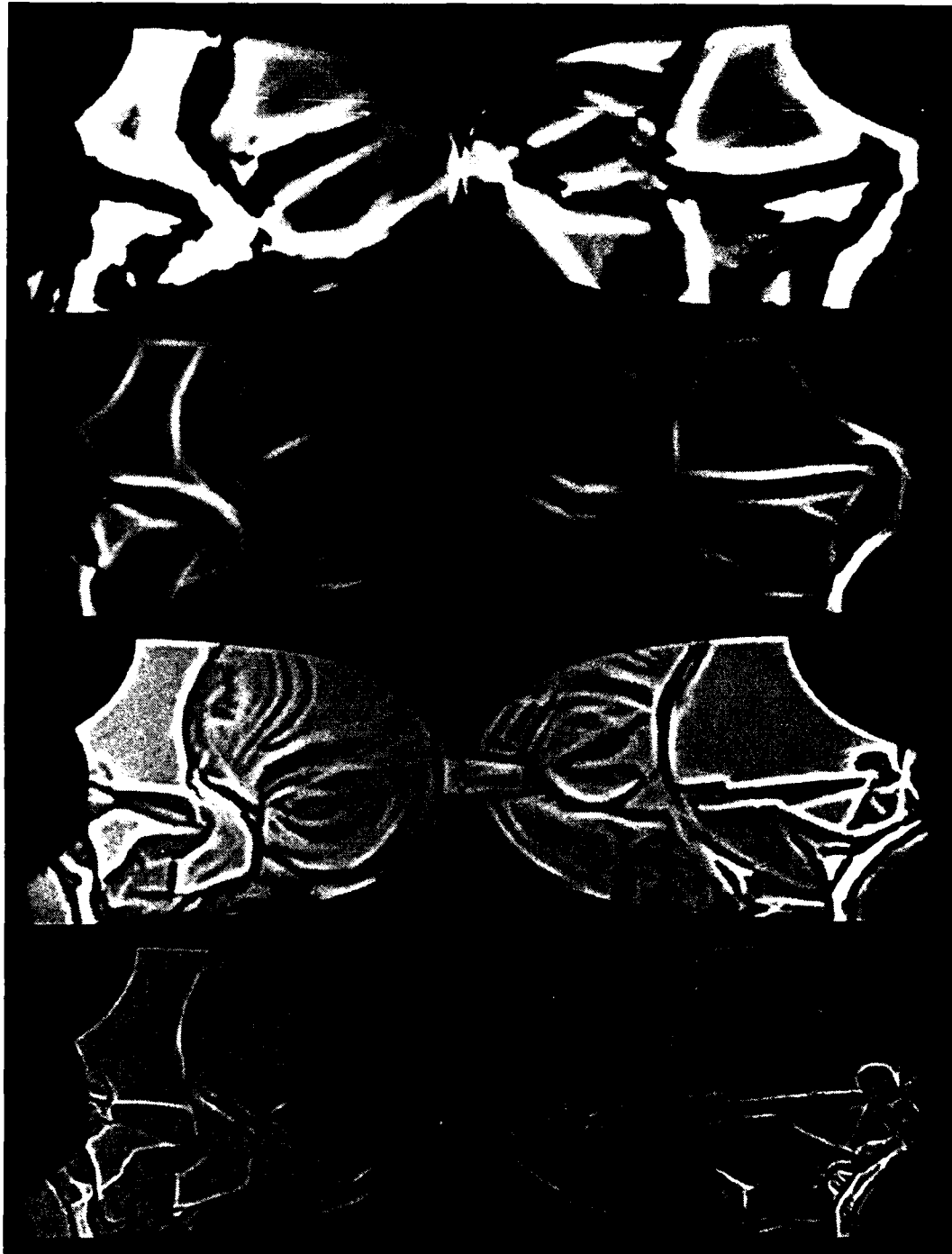


FIGURE 7 CORTICAL PROJECTION OF CONFERENCE ROOM SCENE, FIXATED ON MAN'S SPECTACLES; WINNER-TAKE-ALL ALGORITHM AT FOUR GABOR FREQUENCIES, EACH SEPARATED FROM ITS NEIGHBORS BY ONE OCTAVE



Figure 8 shows the comparable Gabor-filtered outputs for another fixation point of the same scene, located at the young lady's left eye. There is more information about her face in this set of images, but less about other parts of the scene. Again, all four Gabor frequencies seem to contain important information, although the lowest frequencies might be quite indecipherable without the higher ones. The contrast of each of these images has been normalized by the computer output routine to fill the available tone scale, which tends to make all the Gabor frequencies equally visible in each scene.

Figure 9 shows the same scene but with the fixation point moved again, to the center of the notebook. Compare the edges and spine of the notebook with the representation of the two faces. The notebook shows up clearly at all frequencies because it is near the center of the visual field, while the faces out in the periphery are essentially gone at the two lowest frequencies. That also happens subjectively with real retinal images, and this provides an example of the kind of insight that our results can provide.

Effects of this kind have led some investigators to question whether spatial filtering in the periphery is really just a scaled version of foveal filtering--perhaps it is fundamentally different in some way. But the fact that the present model shows similar effects indicates that the scaling of peripheral responses, together with the complex interaction between retinocortical magnification and homogeneous cortical processing, may be sufficient to account for them.

Cortical output images with similar variations of fixation point and Gabor frequency could of course be created for any scene that is available in computer-readable form. Here we provide one further example from our library: the familiar "Lena" portrait, which is widely used in image-coding studies.

Figures 10 through 12 show cortically filtered images with three different fixations of this scene. The four images in each figure represent our standard Gabor frequencies, each one octave from its neighbors. The first fixation point (Figure 10) is on the bridge of the young lady's nose. The second is below and to the left of the first, and the third is above and to the right, both on the brim of her hat (see Figure 9 of Annual Report 2).

In Figure 10, the eyes can be identified at all frequencies except the lowest, but that image is only interpretable with the aid of the higher-frequency information.

In Figure 11 (where the fixation point is on the left brim of the lady's hat), the first and second images from the top both seem completely unintelligible. Only in the third image from the top does her face appear, in the right hemisphere projection. But now it is distorted into an ugly, menacing grimace. Much of the left hemisphere is occupied by the fringe hanging from her hat. It is not a surprise that this resembles the Gabor-filtered images of our white-noise test target (Figure 6).



FIGURE 8 SAME AS FIGURE 7 BUT FIXATED ON WOMAN'S LEFT EYE



FIGURE 9 SAME AS FIGURES 7 AND 8 BUT FIXATED ON CENTER OF NOTEBOOK



FIGURE 10 CORTICAL PROJECTION OF "LENA" PORTRAIT, FIXATED ON BRIDGE OF HER NOSE; SAME COMPUTATION AT SAME FREQUENCIES AS FIGURES 7-9

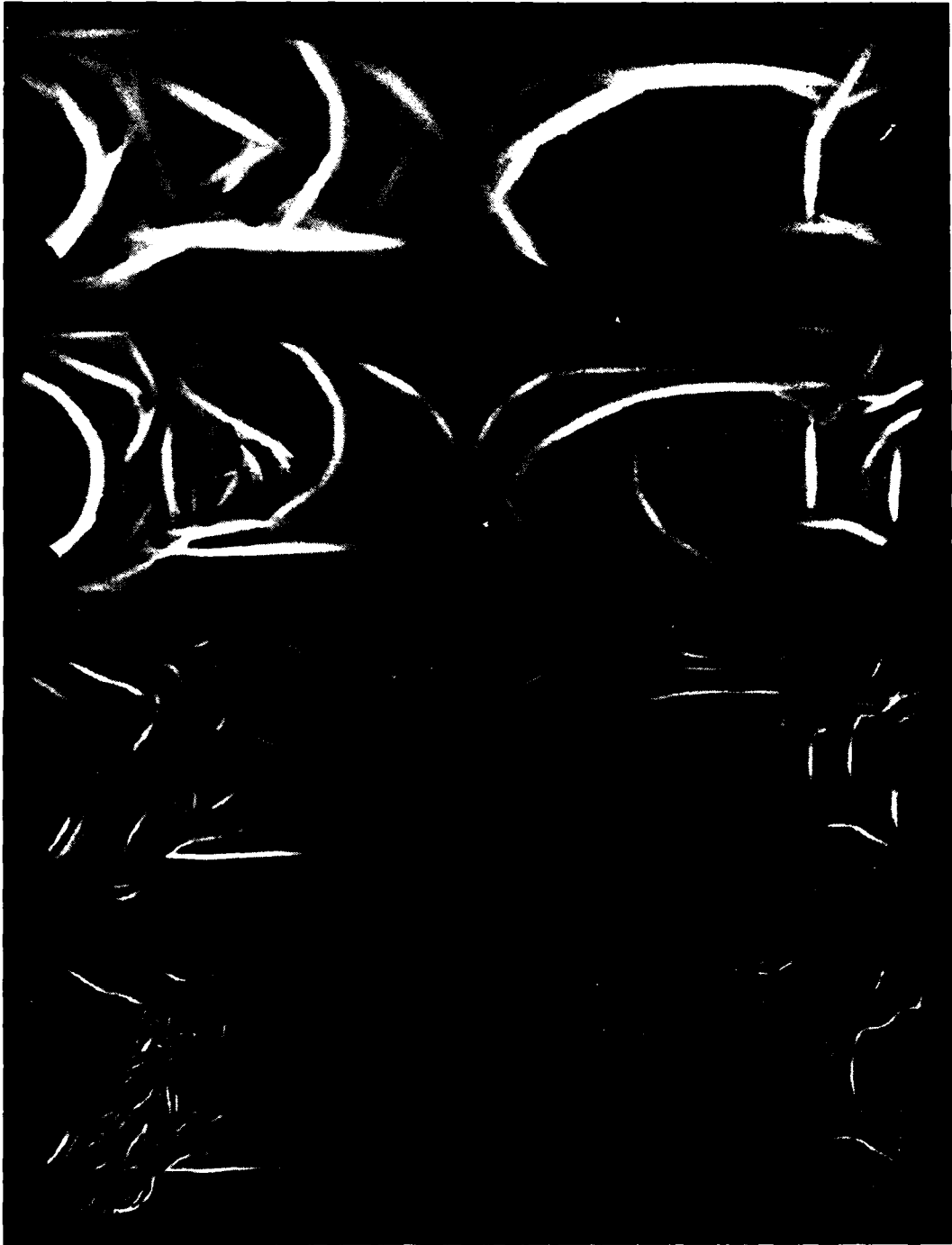


FIGURE 11 SAME AS FIGURE 10 BUT FIXATED ON LEFT BRIM OF MODEL'S HAT



FIGURE 12 SAME AS FIGURES 10 AND 11 BUT FIXATED ON UPPER RIGHT BRIM OF MODEL'S HAT

In Figure 12, the fixation is above and to the right of the other two, so the face now appears in the lower part of the left hemisphere. Its orientation is rotated almost 90 degrees, just as the extrafoveal faces in the conference-room scene were rotated (Figure 9). The face can be made out clearly at the highest Gabor frequency, in the bottom image, but it looks like a quite different person from the one in the previous figure--wan and wistful, instead of mean and nasty. (Of course that doesn't really explain why one looks directly at a person in order to judge her expression!)

#### E. THE STABLE FRAME

We believe that Figures 7 through 12 illustrate the form in which spatial information from the parvo pathways is represented at an early stage of processing in the visual cortex (V1); what happens to it after that is not nearly as well established. At some point, however, this information must be integrated into the so-called stable frame. One of the aims of the present project was to gain some insight into how this occurs. The need for such a process is clear from everyday visual experience.

The visual perception of one's environment (a new room, perhaps, not previously encountered) depends entirely on a small, finite number of fixations. That percept can be acquired even with one eye covered. Normally it is extremely stable, like a map or model of the room. The visual cortex, on the other hand, is hard wired to the retina, so its input consists of a series of distorted images that leap and twist and change with every eye movement, even when they are all from *the very same viewpoint*. This is illustrated by the three fixations shown in Figure 13 (our conference-room scene). Somehow these distorted images must be integrated into a stable framework that forms a unified percept. But how?

Only one Gabor frequency is shown for all three fixations in Figure 13, but that is sufficient to illustrate the difficulty. A full-spectrum image can readily be synthesized from its various Gabor frequencies at the same fixation, but that would be uniquely pointless, since it would merely restore the cortical input. The problem is how to deal with the cortical geometry of *different* fixations (in any or all frequencies) without simply restoring the retinal image, which seems equally pointless.

Indeed, it is not clear that a set of images like those of Figure 13 could be efficiently integrated by any practical procedure. Head and eye position information is presumably available, but that is not sufficient to combine such differently distorted images--we would need the retinal coordinates of every pixel.

Our results suggest that such an image-processing, geometrical solution to this problem is highly improbable. But perhaps no more improbable than the alternative--that these cortical images must be converted into symbolic, cognitive terms before they can be integrated.



FIGURE 13 CORTICAL PROJECTIONS FOR THREE FIXATIONS OF CONFERENCE ROOM SCENE AT THE SAME, INTERMEDIATE GABOR FREQUENCY; SOMEHOW THE INFORMATION IN SEVERAL PROJECTIONS LIKE THIS MUST BE INTEGRATED TO FORM A STABLE PERCEPT OF THE SCENE



### III SIGNIFICANT MILESTONES

Noteworthy accomplishments during the three years of this project have included:

- Exploitation of the hardware and software tools of computer vision (Symbolics 3600 LISP machines) to develop a simulation of human early vision that can display the information it transmits at any stage (as a CRT image).
- Development of a retina-like inhomogeneous spatial filtering algorithm, with a central fixation point whose coordinates can be chosen anywhere in the input image.
- Study of the information contained in a small number of inhomogeneous representations of the same scene with various fixation points.
- Derivation of retinal receptive-field locations by inverse projection from a homogeneous cortical array, using the inverse of the Schwartz (conformal mapping) transformation.
- Modeling of retinal inhomogeneity and retinocortical magnification as independent processes, to facilitate the accurate representation of cortical receptive fields.
- Rapid assembly of locally oriented outputs from orientation-selective units, creating "winner-take-all" cortical images for any given size of (Gabor type) spatial filter.
- Conclusion that the stable frame is probably not achieved by any straightforward image-processing operations that can be performed on information transmitted by the primary visual cortex (V1).

#### IV RECOMMENDATIONS FOR FURTHER WORK

Among other aspects of early vision, chromatic responses and spatiochromatic interactions were intentionally omitted from the first phase of this project, to keep the effort of manageable size. It now appears that this is an important direction in which our cortical modeling could and should be extended.

The chromatic and spatial aspects of early vision are so intimately intertwined that each must be studied in order to fully understand the other. In the recent work of Dr. E. Martinez-Uriegas, of our laboratory, the progression from cone mosaic to retinal, lateral geniculate nucleus (LGN) and finally, cortical receptive fields, emerges as a remarkably efficient coding system. In order to pack the spatial, temporal, and chromatic parameters of the stimulus into the least channel capacity, these inputs are multiplexed in such a way that their familiar sensory correlates are disguised in the optic-nerve and optic-tract signals.

The cortical decoding (demultiplexing) process, as modeled by Martinez-Uriegas, is equally efficient. For example, the wiring required for orientation selectivity (a phenomenon that has been well established physiologically and psychophysically) now emerges as consistent with (and essential to) the process that sorts out opponent-color and luminance responses.

Clearly this kind of detailed modeling can lead to a deeper understanding than we have achieved so far. We urge that further studies along these lines be undertaken.

## **V LIST OF RESEARCH PERSONNEL**

D. H. Kelly, Principal Investigator (Visual Sciences Program)

John D. Peters, Research Assistant (Visual Sciences Program):

Mr. Peters served as chief programmer for this project, writing essentially all the code required to produce the results reported here.

Yvan G. Leclerc, Computer Scientist (Artificial Intelligence Center)

Grahame Smith, Senior Computer Scientist (Artificial Intelligence Center):

Dr. Smith served as our liaison with the Artificial Intelligence Center and its facilities until November 1988, when he moved to the Australian Artificial Intelligence Institute. Since then his role has been performed by Dr. Leclerc.

## VI INTERACTIONS WITH SCIENTIFIC COMMUNITY

By 1989, this work was considered mature enough to be reported in the open literature, and a suitable forum was sought. At the 1990 SPIE/SPSE Symposium on Electronic Imaging Science & Technology, the Principal Investigator was invited to give the opening paper of the session on Visual Models: Spatial Vision and Spatiotemporal Interactions, and he used the occasion to make a report entitled "Retinocortical Processing of Spatial Patterns." Reprints (from the SPIE Proceedings) are available from the author.

More formal (archival) publication may also be undertaken. Y. Y. Zeevi, editor of the recently launched Journal of Visual Communication and Image Representation, has solicited material from the Principal Investigator; a more detailed report on this project may be appropriate for that journal.

## REFERENCES

- Burt, P. J., and E. H. Adelson, 1983: "A multiresolution spline with application to image mosaics," *ACM Trans. on Graphics*, Vol. 2, No. 4, pages 217-236 (October).
- Daugman, J. G., 1980: "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Res.*, Vol. 20, pages 847-856.
- Daugman, J. G., 1983: "Six formal properties of two-dimensional anisotropic visual filters: Structural principles and frequency/orientation selectivity," *IEEE Trans. Syst., Man, Cybern.*, Vol. 13, pages 882-887.
- Daugman, J. G., 1985: "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A.*, Vol. 2, pages 1160-1169.
- Daugman, J. G., 1987: "Image analysis and compact coding by oriented 2D Gabor primitives," *SPIE Proceedings*, Vol. 758 (April).
- Daugman, J. G., 1988: "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE ASSP*, Vol. 36, pages 1169-1179.
- De Valois, R. L., 1971: "Contribution of different lateral geniculate cell types to visual behavior," *Vision Res.*, Supplement No. 3, pages 383-396.
- De Valois, R. L., 1978: "Spatial processing of luminance and color information," *Invest. Ophthalmol. Visual Sci.*, (September (ARVO Symposium Abstracts)), pages 834-835.
- De Valois, R. L., D. G. Albrecht, and L. G. Thorell, 1982: "Spatial frequency selectivity of cells in macaque visual cortex," *Vision Res.*, Vol. 22, pages 545-559.
- De Valois, R. L., and K. K. De Valois, 1980: "Spatial vision," *Ann. Rev. Psychol.*, Vol. 31, pages 309-341.
- De Valois, K. K., R. L. De Valois, and E. W. Yund, 1979: "Responses of striate cortex cells to grating and checkerboard patterns," *J. Physiol.*, Vol. 291, pages 483-505.
- De Valois, R. L., L. G. Thorell, and D. G. Albrecht, 1985: "Periodicity of striate-cortex-cell receptive fields," *J. Opt. Soc. Am. A.*, Vol. 2, pages 1115-1123.
- De Valois, R. L., E. W. Yund, and N. Hepler, 1982: "The orientation and direction selectivity of cells in macaque visual cortex," *Vision Res.*, Vol. 22, pages 531-544.

- Enroth-Cugell, C., and J. G. Robson, 1966: "The contrast sensitivity of retinal ganglion cells of the cat," *J. Physiol.*, Vol. 187, pages 517-552.
- Enroth-Cugell, C., J. G. Robson, D. E. Schweitzer-Tong, and A. B. Watson, 1983: "Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation," *J. Physiol.*, Vol. 341, pages 279-307.
- Fisher, B., 1973: "Overlap of receptive field centers and representation of the visual field in the cat's optic tract," *Vision Res.*, Vol. 13, pages 2113-2120.
- Frishman, L., A. W. Freeman, J. B. Troy, D. E. Schweitzer-Tong, and C. Enroth-Cugell, 1987: "Spatiotemporal frequency responses of cat retinal ganglion cells," *J. Gen. Physiol.*, Vol. 89, pages 599-628.
- Jones, J., and L. Palmer, 1987: "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, Vol. 58, pages 1233-1258.
- Kelly, D. H., 1975: "Spatial frequency selectivity in the retina," *Vision Res.*, Vol. 15, pages 665-672.
- Kelly, D. H., 1979: "Motion and vision. II. Stabilized spatio-temporal threshold surface," *J. Opt. Soc. Am.*, Vol. 69, pages 1340-1349.
- Kelly, D. H., and C. A. Burbeck, 1984: "Critical problems in spatial vision," *CRC Cr. Rev. Biomed. Engng.*, Vol. 10, pages 125-177.
- Kelly, D. H., 1985: "Retinal inhomogeneity. III. Circular-retina theory," *J. Opt. Soc. Am. A.*, Vol. 2, pages 810-819.
- Kelly, D. H., and A. P. Pentland, 1985: "Why we see the whole world sharply: Eye movements, retinal inhomogeneity and cortical processing," unpublished report, SRI International.
- Kelly, D. H., 1990: "Retinocortical processing of spatial patterns," *Proc. Soc. Photo-Optical Engineers*, Vol. 1249.
- Letelier, J. C., and F. Varela, 1984: "Why the cortical magnification factor in rhesus is isotropic," *Vision Res.*, Vol. 24, pages 1091-1095.
- Linsenmeier, R. A., L. J. Frishman, H. G. Jakiela, and C. Enroth-Cugell, 1982: "Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements," *Vision Res.*, Vol. 22, pages 1173-1183.
- Marr, D., and E. Hildreth, 1980: "Theory of edge detection," *Proc. R. Soc. London, Ser. B* 290, pages 199-218.
- Porat, M., and Y. Y. Zeevi, 1988: "The generalized Gabor scheme of image representation in biological and machine vision," *Trans. IEEE PAMI*, Vol. 10, No. 4, pages 452-468.

- Rovamo, J., 1978: "Receptive field density of retinal ganglion cells and cortical magnification factor in man," *Medical Biology*, Vol. 56, pages 97-102.
- Rovamo, J., and V. Virsu, 1979: "An estimation and application of the human cortical magnification factor," *Exp. Brain Res.*, Vol. 37, pages 495-510.
- Sakitt, B., and H. Barlow, 1982: "A model for the economical encoding of the visual image in the cerebral cortex," *Biol. Cybern.*, Vol. 43, pages 97-108.
- Schwartz, E. L., 1977: "Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception," *Biol. Cybern.*, Vol. 25, pages 97-108.
- Schwartz, E. L., 1980: "Computational anatomy and functional architecture of striate cortex," *Vision Res.*, Vol. 20, pages 645-669.
- Schwartz, E., 1985: "On the mathematical structure of the visuotopic mapping of macaque striate cortex," *Science*, Vol. 227, pages 1065-1066.
- Schwartz, E. L., and B. Merker, 1986: "Computer-aided neuroanatomy: Differential geometry of cortical surfaces and an optimal flattening algorithm," *IEEE CG&A.*, Vol. 6, pages 36-44.
- Thorell, L. G., R. L. De Valois, and D. G. Albrecht, 1984: "Spatial mapping of monkey V1 cells with pure color and luminance stimuli," *Vision Res.*, Vol. 24, No. 7, pages 751-769.
- Tootell, R. B. H., M. S. Silverman, and R. L. De Valois, 1981: "Spatial frequency columns in primary visual cortex," *Science*, Vol. 214, pages 813-815 (13 November).
- Tootell, R. B. H., M. S. Silverman, E. Switkes, and R. L. De Valois, 1982: "Deoxyglucose analysis of retinotopic organization in primate striate cortex," *Science*, Vol. 218, pages 902-904.
- Tootell, R. B. H., M. S. Silverman, R. L. De Valois, and G. H. Jacobs, 1983: "Functional organization of the second cortical visual area in primates," *Science*, Vol. 220, pages 737-739, (13 May).
- Tootell, R. B. H., S. L. Hamilton, M. S. Silverman, and E. Switkes, 1988: "Functional anatomy of macaque striate cortex. I. Ocular dominance, binocular interactions, and baseline conditions," *J. Neuroscience*, Vol. 8, No. 5, pages 1500-1530.
- Tootell, R. B. H., E. Switkes, M. S. Silverman, and S. L. Hamilton, 1988: "Functional anatomy of macaque striate cortex. II. Retinotopic organization" *J. Neuroscience*, Vol. 8, pages 1531-1568.
- Tootell, R. B. H., M. S. Silverman, S. L. Hamilton, R. L. De Valois, and E. Switkes, 1988: "Functional anatomy of macaque striate cortex. III. Color," *J. Neuroscience*, Vol. 8, No. 5, pages 1569-1593.

- Tootell, R. B. H., S. L. Hamilton, and E. Switkes, 1988: "Functional anatomy of macaque striate cortex. IV. Contrast and magno-parvo streams," *J. Neuroscience*, Vol. 8, No. 5, pages 1594-1609.
- Tootell, R. B. H., M. S. Silverman, S. L. Hamilton, E. Switkes, and R. L. De Valois, 1988: "Functional anatomy of macaque striate cortex. V. Spatial frequency," *J. Neuroscience*, Vol. 8, No. 5, pages 1610-1624.
- VanDoorn, A. J., J. J. Koenderink, and M. A. Bouman, 1972: "The influence of retinal inhomogeneity on the perception of spatial patterns," *Kybernetik*, Vol. 10, pages 223-230.
- Webster, M. A., and R. L. De Valois, 1985: "Relationship between spatial-frequency and orientation tuning of striate-cortex cells," *J. Opt. Soc. Am. A.*, Vol. 2, No. 7, pages 1124-1132.
- Wilson, J. R., and S. M. Sherman, 1976: "Receptive-field characteristics of neurons in cat striate cortex; changes with visual field eccentricity," *J. Neurophysiol.*, Vol. 39, pages 512-533.